PERGAMON Computers and Mathematics with Applications 38 (1999) 31–39

An International Journal
computers & mathematics with applications

www.elsevier.nl/locate/camwa

Asymptotic Properties of an Estimator in Errors-In-Variables Models in the Presence of Validation Data

I. FAZEKAS[1]
Institute of Mathematics and Informatics, University of Debrecen
H-4010 Debrecen, P.O. Box 12, Hungary
fazekasi@math.klte.hu

S. BARAN[2]
Institute of Mathematics and Informatics, University of Debrecen
H-4010 Debrecen, P.O.Box 12, Hungary
barans@math.klte.hu

J. LAURIDSEN
Department of Statistics and Demography, Odense University
Campusvej 55, DK-5230 Odense M, Denmark
jtl@busieco.ou.dk

(Received and accepted March 1999)

Abstract—Structural errors-in-variables models with dependent spatial observations are studied. The presence of validation data is assumed. An estimator for regression parameters proposed by Lee and Sepanski [1] is studied. Consistency and asymptotic normality of the estimator are established in the case of increasing domain. Infill asymptotic properties are described. Simulation results are also presented. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords—Measurement errors, Errors-in-variables, Validation, Mixing property, Consistency of estimators, Random fields.

1. INTRODUCTION

Analyzing economics data, one often encounters nonstandard statistical problems. This paper deals with three problems: measurement errors, dependent observations, and spatial data. Consider the model

$$y_i = g(\mathbf{x}_i, \beta_0) + \varepsilon_i, \tag{1.1}$$

where y_i is observed, ε_i is random error term, $\mathbf{x}_i$ is the k-dimensional column vector of explanatory variables, $i \in T$, g is a known function, and β_0 is the true value of the unknown l-dimensional parameter β to be estimated. Here, T is the set of points where the random field is defined. Explanatory variables $\mathbf{x}_i$ are measured with error

$$\tilde{\mathbf{x}}_i \text{ measures } \mathbf{x}_i \text{ with error.} \tag{1.2}$$

[1]The research was accomplished while this author visited Odense University, Denmark.
[2]Supported by Hungarian Ministry of Culture and Education under Grant No. 230/1996.

PII:S0898-1221(99)00212-6

Models having the form (1.1),(1.2) are called errors-in-variables regression models. Errors-in-variables models have been considered for about half a century. A summary of results is given in Fuller's textbook [2]. In this paper, discussion will be confined to the structural case, i.e., when $\mathbf{x}_i$s are random variables.

In [1], a consistent procedure assuming validation data and based on least squares methods was proposed. In that paper, i.i.d. data were considered. The aim of the present paper is to extend the results of [1] to dependent data. We concentrate on the case where measurement errors are only in explanatory variables $\mathbf{x}_i$. We assume that our data are observations of an underlying random field. We suppose that the field satisfies weak dependence conditions. We apply α-mixing conditions but one can find similar results using other weak dependence conditions.

The model and the estimator are described in Section 2. In Section 3, simulation results are presented showing the advantage of the Lee-Sepanski estimator. Consistency and asymptotic normality are proved in Section 4. The proofs given here are versions of those in [1]. In Section 5, it is shown that the estimator is not consistent for dense observations in a fixed domain. Note that a detailed mathematical analysis of this topic can be found in [3].

The following notation is used. $\mathbf{Z}^d$ is the set of d-dimensional lattice points, $\mathbf{R}$ is the real line, $\mathbf{R}^p$ is the p-dimensional Euclidean space with norm $\| \ \|$. Superscript $\top$ denotes the transpose of a matrix. Vectors without $\top$ are column vectors. I denotes the identity matrix. We shall denote different constants with the same letter c. $|\mathcal{P}|$ denotes the cardinality of the finite set $\mathcal{P}$. We shall suppose the existence of an underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$. $\mathbf{E}$ stands for the expectation.

$$\|\eta\|_p = \{\mathbf{E}\|\eta\|^p\}^{1/p}, \qquad 1 \le p < \infty,$$

is the norm in L_p. By $o_P(1)$, a quantity converging to zero in probability is denoted. Sign $\Rightarrow$ denotes convergence in distribution. $\mathcal{N}(m, \Sigma)$ stands for the (vector) normal distribution with mean (vector) m and covariance (matrix) Σ.

2. THE MODEL AND THE ESTIMATOR

Let us consider the model (1.1),(1.2).

We have *primary data* that can be considered as observations of the random fields y_i, $\tilde{\mathbf{x}}_i$, $i \in T$, at certain locations. They are denoted by $(y_{\circ i}, \tilde{\mathbf{x}}_{\circ i})$, $i \in \mathcal{P}_n$. Here $\mathcal{P}_n \subset T$ is a finite set for each $n = 1, 2, \dots$. Let $\mathbf{Y}_\circ = (y_{\circ 1}, \dots, y_{\circ |\mathcal{P}_n|})^\top$ be $|\mathcal{P}_n| \times 1$ and $\tilde{\mathbf{X}}_\circ = (\tilde{\mathbf{x}}_{\circ 1}, \dots, \tilde{\mathbf{x}}_{\circ |\mathcal{P}_n|})^\top$ be $|\mathcal{P}_n| \times k$ matrices of the primary data. Assume that a known k_1-dimensional vector-valued function $\mathbf{z}_i$ of $\mathbf{x}_i$ is available for which $g(\mathbf{x}_i, \beta_0) \approx \mathbf{z}_i^\top \gamma(\beta_0)$, i.e., $g(\mathbf{x}_i, \beta_0)$ can be approximated with a linear function of $\mathbf{z}_i$. Let $\mathbf{Z}_\circ = (\mathbf{z}_{\circ 1}, \dots, \mathbf{z}_{\circ |\mathcal{P}_n|})^\top$ be the $|\mathcal{P}_n| \times k_1$ matrix of $\mathbf{z}_i$s based on the primary data. To find asymptotic behaviour of the estimator, we shall suppose that $\mathcal{P}_n$, $n = 1, 2, \dots$, is an increasing sequence of finite subsets of T with $|\mathcal{P}_n| \to \infty$, as $n \to \infty$. Denote by $\mathcal{P}_\infty$ the set $\cup_{n=1}^\infty \mathcal{P}_n$.

Assume that for explanatory variables, *validation data* are available. Validation data can be considered as observations on random fields $\mathbf{x}_i$, $\tilde{\mathbf{x}}_i$, $i \in T$, taken at some points which set of points is not equal to $\mathcal{P}_n$. Validation data are denoted by $(\mathbf{x}_{\vee i}, \tilde{\mathbf{x}}_{\vee i})$, where $\mathbf{x}_{\vee i}$, is a measurement without error, while $\tilde{\mathbf{x}}_{\vee i}$ is the corresponding measurement with error, $i \in \mathcal{V}_m$. Here $\mathcal{V}_m \subset T$ is a finite set for each $m = 1, 2, \dots$. We need the values of g at the precise validation data: $g(\mathbf{X}_\vee, \beta) = (g(\mathbf{x}_{\vee 1}, \beta), \dots, g(\mathbf{x}_{\vee |\mathcal{V}_m|}, \beta))^\top$ (a matrix of type $|\mathcal{V}_m| \times 1$). We also need the values of $\mathbf{z}_i$ at the validation data with error: $\mathbf{Z}_\vee = (\mathbf{z}_{\vee 1}, \dots, \mathbf{z}_{\vee |\mathcal{V}_m|})^\top$ (a matrix of type $|\mathcal{V}_m| \times k_1$). To find asymptotic behaviour of the estimator, we shall suppose that $\mathcal{V}_m$, $m = 1, 2, \dots$, is an increasing sequence of finite subsets of T with $|\mathcal{V}_m| \to \infty$, as $m \to \infty$. Denote by $\mathcal{V}_\infty$ the set $\cup_{m=1}^\infty \mathcal{V}_m$.

The estimator $\hat{\beta} = \hat{\beta}_{n,m}$ proposed in [1] is the minimum point of

$$Q_{n,m}(\beta) = \frac{1}{|\mathcal{P}_n|} \left\| \mathbf{Y}_\circ - \mathbf{Z}_\circ \left(\mathbf{Z}_\vee^\top \mathbf{Z}_\vee \right)^{-1} \mathbf{Z}_\vee^\top g(\mathbf{X}_\vee, \beta) \right\|^2. \tag{2.1}$$

The intuitive background of the above estimator is the following. First we approximate $g(\mathbf{x}_i, \beta)$ with a function of $\tilde{\mathbf{x}}_i$. To this end, we use the linear model $g(\mathbf{x}_i, \beta) \approx \mathbf{z}_i^\top \gamma(\beta)$ and validation data. The least squares estimator in the linear model gives that the estimator of $\gamma(\beta)$ is $(\mathbf{Z}_V^\top \mathbf{Z}_V)^{-1} \mathbf{Z}_V^\top g(\mathbf{X}_V, \beta)$. We replace $\gamma(\beta)$ by this estimator in the linear model which approximates our original one. Then, we use primary data and least squares estimator to obtain $\hat{\beta}$.

3. SIMULATION RESULTS

The examples below show the behaviour of the Lee-Sepanski estimator for dependent spatial observations. We compare this estimator with the naive one which is the least squares estimator considering $\tilde{\mathbf{x}}_i$ as it would be an observation without error. The advantage of the Lee-Sepanski estimator is obvious.

Simulations were performed by the help of MATLAB. For minimization, procedure `leastsq` was used (see [4]). We remark that the initial point of the minimization procedure was found using a search in the positive quarter.

EXAMPLE 3.1. LINEAR MODEL. Consider the model

$$y_{i,j} = \theta_0 + \theta_1 x_{i,j} + \varepsilon_{i,j}, \qquad i, j \in T, \tag{3.1}$$

where $T \subset \mathbf{N}^2$ and all variables are one dimensional. The error terms $\varepsilon_{i,j}$ form a stationary Gaussian first-order autoregressive (AR(1)) process (if j is fixed):

$$\varepsilon_{i+1,j} = a\varepsilon_{i,j} + \delta_{i+1,j}, \tag{3.2}$$

for every i, j, where variables $\delta_{i,j}$ are independent standard normal and the initial distribution of $\varepsilon_{1,j}$ is chosen according stationarity. The value of a is fixed on 0.2. Explanatory variables $x_{i,j}$ form a spatial moving average (MA) process:

$$x_{i,j} = \frac{1}{9} \sum_{|u-i| \le 1, |v-j| \le 1} \eta_{u,v}, \tag{3.3}$$

for every i, j, where $\eta_{u,v}$ are independent standard normal. The observable explanatory variables $\tilde{x}_{i,j}$ are

$$\tilde{x}_{i,j} = x_{i,j} + \xi_{i,j}, \tag{3.4}$$

where the errors $\xi_{i,j}$ are i.i.d. normal. To find the Lee-Sepanski estimator,

$$\mathbf{z}_{i,j} = (1, x_{i,j})^\top \tag{3.5}$$

was chosen for each i, j.

In four special cases, the mean and the variance of the naive and the Lee-Sepanski estimator were calculated. In each case, 2000 replications were performed.

CASE 1. Primary data are taken at locations $\mathcal{P} = \{1 \le i \le 20,\ 1 \le j \le 20\}$, validation data are taken at locations $\mathcal{V} = \{1 \le i \le 30,\ 1 \le j \le 7\}$, $\xi_{i,j} \sim \mathcal{N}(0, 0.5)$. For $\theta_0 = 0.5$ and $\theta_1 = 1$, we have for the naive estimator: mean$(\tilde{\theta}_0) = 0.5025$, var$(\tilde{\theta}_0) = 0.0050$, mean$(\tilde{\theta}_1) = 0.3038$, var$(\tilde{\theta}_1) = 0.0092$, and for the Lee-Sepanski estimator: mean$(\hat{\theta}_0) = 0.5031$, var$(\hat{\theta}_0) = 0.0063$, mean$(\hat{\theta}_1) = 1.0312$, var$(\hat{\theta}_1) = 0.1268$.

CASE 2. Primary data are taken at locations $\mathcal{P} = \{1 \le i \le 27,\ 1 \le j \le 30\}$, validation data are taken at locations $\mathcal{V} = \{1 \le i \le 40,\ 1 \le j \le 10\}$, $\xi_{i,j} \sim \mathcal{N}(0, 0.75)$. For $\theta_0 = 0.5$ and $\theta_1 = 1$, we have for the naive estimator: mean$(\tilde{\theta}_0) = 0.4993$, var$(\tilde{\theta}_0) = 0.0028$, mean$(\tilde{\theta}_1) = 0.1642$, var$(\tilde{\theta}_1) = 0.0023$, and for the Lee-Sepanski estimator: mean$(\hat{\theta}_0) = 0.4990$, var$(\hat{\theta}_0) = 0.0035$, mean$(\hat{\theta}_1) = 1.0275$, var$(\hat{\theta}_1) = 0.1055$.

CASE 3. The only difference from Case 2 is that the errors in explanatory variables are different for primary data and for validation data. Namely for validation data $\tilde{x}_{i,j} = x_{i,j} + \zeta_{i,j}$, where $\zeta_{i,j}$ are i.i.d. with $\zeta_{i,j} \sim \mathcal{N}(0, 0.75)$. For $\theta_0 = 0.5$ and $\theta_1 = 1$, we have for the naive estimator: $\mathrm{mean}(\tilde{\theta}_0) = 0.4990$, $\mathrm{var}(\tilde{\theta}_0) = 0.0020$, $\mathrm{mean}(\tilde{\theta}_1) = 0.1631$, $\mathrm{var}(\tilde{\theta}_1) = 0.0020$, and for the Lee-Sepanski estimator: $\mathrm{mean}(\hat{\theta}_0) = 0.4993$, $\mathrm{var}(\hat{\theta}_0) = 0.0035$, $\mathrm{mean}(\hat{\theta}_1) = 1.0228$, $\mathrm{var}(\hat{\theta}_1) = 0.1136$.

CASE 4. The only difference from Case 3 is that primary and validation data are taken from disjoint regions. Primary data are taken at locations $\mathcal{P} = \{1 \le i \le 27,\ 1 \le j \le 30\}$, validation data are taken at locations $\mathcal{V} = \{28 \le i \le 40,\ 1 \le j \le 30\}$. For $\theta_0 = 0.5$ and $\theta_1 = 1$, we have for the naive estimator: $\mathrm{mean}(\tilde{\theta}_0) = 0.4990$, $\mathrm{var}(\tilde{\theta}_0) = 0.0020$, $\mathrm{mean}(\tilde{\theta}_1) = 0.1631$, $\mathrm{var}(\tilde{\theta}_1) = 0.0024$, and for the Lee-Sepanski estimator: $\mathrm{mean}(\hat{\theta}_0) = 0.4985$, $\mathrm{var}(\hat{\theta}_0) = 0.0047$, $\mathrm{mean}(\hat{\theta}_1) = 1.0292$, $\mathrm{var}(\hat{\theta}_1) = 0.1239$.

EXAMPLE 3.2. QUADRATIC MODEL. Consider the model

$$y_{i,j} = \left(\theta^\top \mathbf{x}_{i,j}\right)^2 + \varepsilon_{i,j}, \qquad i,j \in T, \tag{3.6}$$

where $T \subset \mathbf{N}^2$, $y_{i,j}$ and $\varepsilon_{i,j}$ are one dimensional, $\theta^\top = (\theta_0, \theta_1)$ and $\mathbf{x}_{i,j}^\top = (x_{i,j}^{(1)}, x_{i,j}^{(2)})$ are two dimensional. The error term $\varepsilon_{i,j}$ is the same stationary Gaussian AR(1) process as in Example 3.1. Both $x_{i,j}^{(1)}$ and $x_{i,j}^{(2)}$, $i,j \in T$, form spatial MA processes of form (3.3) and these processes are independent. The observable explanatory variables $\tilde{x}_{i,j}^{(h)}$ are

$$\tilde{x}_{i,j}^{(h)} = x_{i,j}^{(h)} + \xi_{i,j}^{(h)}, \tag{3.7}$$

where the errors $\xi_{i,j}^{(h)}$ are i.i.d. with $\xi_{i,j}^{(h)} \sim \mathcal{N}(0, 0.5)$, $i,j \in T$, $h = 1,2$. To find the Lee-Sepanski estimator

$$\mathbf{z}_{i,j} = \left(\left(x_{i,j}^{(1)}\right)^2, \left(x_{i,j}^{(2)}\right)^2, x_{i,j}^{(1)} x_{i,j}^{(2)}\right)^\top, \tag{3.8}$$

was chosen for each $i,j \in T$.

Primary data are taken at locations $\mathcal{P} = \{1 \le i \le 27,\ 1 \le j \le 30\}$, validation data are taken at locations $\mathcal{V} = \{1 \le i \le 40,\ 1 \le j \le 10\}$. 2000 replications were performed and the mean and the variance of the naive and the Lee-Sepanski estimator were calculated. For $\theta_0 = 0.5$ and $\theta_1 = 1$ we have for the naive estimator: $\mathrm{mean}(\tilde{\theta}_0) = 0.1370$, $\mathrm{var}(\tilde{\theta}_0) = 0.0140$, $\mathrm{mean}(\tilde{\theta}_1) = 0.2133$, $\mathrm{var}(\tilde{\theta}_1) = 0.0094$, and for the Lee-Sepanski estimator: $\mathrm{mean}(\hat{\theta}_0) = 0.5105$, $\mathrm{var}(\hat{\theta}_0) = 0.1855$, $\mathrm{mean}(\hat{\theta}_1) = 0.8463$, $\mathrm{var}(\hat{\theta}_1) = 0.1522$.

Simulation evidence of the asymptotic normality can be found in [3].

4. CONSISTENCY AND ASYMPTOTIC NORMALITY FOR INCREASING DOMAIN

We shall use the notion of α-mixing for random fields in the sense of [5]. Let η_t, $t \in \mathbf{Z}^d$, be a random field. For $\Lambda \subseteq \mathbf{Z}^d$, let $\mathcal{A}_\Lambda$ denote the σ-algebra generated by $\{\eta_t : t \in \Lambda\}$. Let ϱ denote the distance induced by the max norm in $\mathbf{Z}^d$: $\varrho(\mathbf{u}, \mathbf{v}) = \max_{1 \le i \le d} |u_i - v_i|$, for $\mathbf{u} = (u_1, \dots, u_d)^\top$, $\mathbf{v} = (v_1, \dots, v_d)^\top \in \mathbf{Z}^d$. Denote by $\varrho(\Lambda_1, \Lambda_2)$ the distance of Λ_1 and Λ_2 for $\Lambda_1,\ \Lambda_2 \subseteq \mathbf{Z}^d$. Let

$$\begin{aligned} &\alpha_\eta(n; u, v) \\ &\quad = \sup\{|\mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathcal{A}_{\Lambda_1},\ B \in \mathcal{A}_{\Lambda_2},\ |\Lambda_1| \le u,\ |\Lambda_2| \le v,\ \varrho(\Lambda_1, \Lambda_2) \ge n\}. \end{aligned}$$

The following lemma can be proved by the method of [5, pp. 25–30].

LEMMA 4.1. *Let η_t, $t \in \mathbf{Z}^d$, be a centered random field, $1 < \tau \le 2$, $\varepsilon > 0$ and suppose that condition*

$$\sum_{r=1}^{\infty} r^{d-1} [\alpha_\eta(r; 1, 1)]^{\varepsilon/(2+\varepsilon)} < \infty \tag{4.1}$$

is satisfied. Let $\mathcal{P}$ be a finite subset of $\mathbf{Z}^d$ and assume that $\mathbf{E}|\eta_t|^{\tau+\varepsilon}$ is finite for every $t \in \mathcal{P}$. Then, there exists a constant K depending only on τ and mixing coefficients $\alpha_\eta(r;1,1)$ such that

$$\mathbf{E}\left|\sum_{t\in\mathcal{P}} \eta_t\right|^\tau \le K \sum_{t\in\mathcal{P}} \left(\mathbf{E}\left|\eta_t\right|^{\tau+\varepsilon}\right)^{\tau/(\tau+\varepsilon)}. \tag{4.2}$$

We remark that inequality (4.2) is valid for (finite-dimensional) vector valued fields η_t, too. In this case, constant K may depend on the dimension of η_t.

In our proofs, we will often use the uniform law of large numbers (ULLN) for the mixing case which is an obvious generalization of the one in [6, p. 116] (see [3,7]).

In this section, the notation is the same as in Section 2.

ASSUMPTIONS.

(A1) The parameter set $\Theta \subset \mathbf{R}^l$ is compact and convex. The true parameter β_0 is in the interior of Θ.
(A2) $g(\mathbf{x},\beta)$ is measurable in $\mathbf{x}$ for each β and it is differentiable in β for each $\mathbf{x}$.
(A3) For the primary data: $(y_{\mathrm{o}t}, \mathbf{z}_{\mathrm{o}t})$ is identically distributed as $(y_\mathrm{o}, \mathbf{z}_\mathrm{o})$ for each $t \in \mathcal{P}_\infty$. Condition (4.1) is satisfied for $\eta_t = (y_{\mathrm{o}t}, \mathbf{z}_{\mathrm{o}t})$, $t \in \mathcal{P}_\infty$.
(A4) For a $\delta > 0$, $\mathbf{E}|y_\mathrm{o}|^{2+\delta}$ and $\mathbf{E}\|\mathbf{z}_\mathrm{o}\|^{2+\delta}$ are finite. $\mathbf{E}(\mathbf{z}_\mathrm{o}\mathbf{z}_\mathrm{o}^\top)$ is nonsingular.
(A5) For the validation data: $(g(\mathbf{x}_{\mathrm{V}t},\beta), \mathbf{z}_{\mathrm{V}t})$ is identically distributed as $(g(\mathbf{x}_\mathrm{V},\beta), \mathbf{z}_\mathrm{V})$ for each $t \in \mathcal{V}_\infty$. Condition (4.1) is satisfied for $\eta_t = (g(\mathbf{x}_{\mathrm{V}t},\beta), \mathbf{z}_{\mathrm{V}t})$, $t \in \mathcal{V}_\infty$.
(A6) For a $\delta > 0$, $\mathbf{E}\|\mathbf{z}_\mathrm{V}\|^{2+\delta}$, and for every $\beta \in \Theta$ $\mathbf{E}|g(\mathbf{x}_\mathrm{V},\beta)|^{2+\delta}$ are finite. $\mathbf{E}(\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top)$ is nonsingular.
(A7) $(\mathbf{E}\mathbf{z}_\mathrm{o}\mathbf{z}_\mathrm{o}^\top)^{-1}\mathbf{E}(\mathbf{z}_\mathrm{o}y_\mathrm{o}) = (\mathbf{E}\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top)^{-1}\mathbf{E}(\mathbf{z}_\mathrm{V}g(\mathbf{x}_\mathrm{V},\beta_0))$.
(A8) $\mathbf{E}[\mathbf{z}_\mathrm{V}(g(\mathbf{x}_\mathrm{V},\beta) - g(\mathbf{x}_\mathrm{V},\beta_0))] \ne 0$ for $\beta \ne \beta_0$.

The following consistency result is an extension of Proposition 2.1 of [1] to mixing random fields.

THEOREM 4.2. *Assume that Conditions (A1)–(A8) are satisfied. Suppose that*

$$\mathbf{E}\left(\sup_{\beta\in\Theta} \|\mathbf{z}_\mathrm{V} g(\mathbf{x}_\mathrm{V},\beta)\|\right) < \infty. \tag{4.3}$$

Then, $\hat\beta$, given by (2.1), is a consistent estimator of β_0, as $n, m \to \infty$.

PROOF. The sketch of the proof (for details, see [1,3]). We have

$$\begin{aligned}
Q_{n,m}(\beta) &= \frac{1}{|\mathcal{P}_n|}\|\mathbf{Y}_\mathrm{o}\|^2 - \frac{2}{|\mathcal{P}_n|}\left(\mathbf{Y}_\mathrm{o}^\top \mathbf{Z}_\mathrm{o}\right)\left(\frac{1}{|\mathcal{V}_m|}\mathbf{Z}_\mathrm{V}^\top\mathbf{Z}_\mathrm{V}\right)^{-1}\left(\frac{1}{|\mathcal{V}_m|}\mathbf{Z}_\mathrm{V}^\top g(\mathbf{X}_\mathrm{V},\beta)\right)\\
&+\frac{1}{|\mathcal{P}_n|}\|\mathbf{Z}_\mathrm{o}\left(\mathbf{Z}_\mathrm{V}^\top\mathbf{Z}_\mathrm{V}\right)^{-1}\mathbf{Z}_\mathrm{V}^\top g(\mathbf{X}_\mathrm{V},\beta)\|^2 \longrightarrow \mathbf{E}y_\mathrm{o}^2 - 2\mathbf{E}\left(y_\mathrm{o}\mathbf{z}_\mathrm{o}^\top\right)\left(\mathbf{E}\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top\right)^{-1}\mathbf{E}\left(\mathbf{z}_\mathrm{V} g(\mathbf{x}_\mathrm{V},\beta)\right)\\
&+\mathbf{E}\left(\mathbf{z}_\mathrm{V}^\top g(\mathbf{x}_\mathrm{V},\beta)\right)\left(\mathbf{E}\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top\right)^{-1}\mathbf{E}\left(\mathbf{z}_\mathrm{o}\mathbf{z}_\mathrm{o}^\top\right)\left(\mathbf{E}\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top\right)^{-1}\mathbf{E}\left(\mathbf{z}_\mathrm{V} g(\mathbf{x}_\mathrm{V},\beta)\right) = Q_\infty(\beta),
\end{aligned} \tag{4.4}$$

in probability, as $n, m \to \infty$. This fact follows from Chebyshev's inequality using Lemma 4.1. By assumption (4.3) and the ULLN, the convergence is uniform in β. Assumptions (A7) and (A8) imply that β_0 is the unique minimum point of $Q_\infty(\beta)$. Hence, $\hat\beta \to \beta_0$ in probability. ∎

ADDITIONAL ASSUMPTIONS.

(A2′) $g(\mathbf{x},\beta)$ is measurable in $\mathbf{x}$ for each β and it is twice continuously differentiable in β for each $\mathbf{x}$.
(A4′) For a $\delta > 0$, $\mathbf{E}|y_\mathrm{o}|^{4+\delta}$ and $\mathbf{E}\|\mathbf{z}_\mathrm{o}\|^{4+\delta}$ are finite. $\mathbf{E}(\mathbf{z}_\mathrm{o}\mathbf{z}_\mathrm{o}^\top)$ is nonsingular.
(A6′) For a $\delta > 0$, $\mathbf{E}\|\mathbf{z}_\mathrm{V}\|^{4+\delta}$, and for every $\beta \in \Theta$ $\mathbf{E}|g(\mathbf{x}_\mathrm{V},\beta)|^{4+\delta}$ are finite. $\mathbf{E}(\mathbf{z}_\mathrm{V}\mathbf{z}_\mathrm{V}^\top)$ is nonsingular.

The following theorem is an extension of Proposition 2.2 of [1] to mixing random fields.

THEOREM 4.3. *Suppose that conditions (A1), (A2′), (A3), (A4′), (A5), (A6′), (A7), (A8), and (4.3) are satisfied. In addition, assume*

$$\mathbf{E}\left(\sup_{\beta\in\Theta}\left\|\frac{\partial g(\mathbf{x}_\vee,\beta)}{\partial\beta}\mathbf{z}_\vee^\top\right\|\right)<\infty, \tag{4.5}$$

$$\mathbf{E}\left(\sup_{\beta\in\Theta}\left\|\frac{\partial^2 g(\mathbf{x}_\vee,\beta)}{\partial\beta\partial\beta_j}\mathbf{z}_\vee^\top\right\|\right)<\infty, \qquad j=1,2,\dots,l, \tag{4.6}$$

$$\mathbf{E}\left\{\frac{\partial g(\mathbf{x}_\vee,\beta_0)}{\partial\beta}\mathbf{z}_\vee^\top\right\} \text{ has rank } l, \tag{4.7}$$

$$\lambda=\lim_{n,m\to\infty}\sqrt{\frac{|\mathcal{P}_n|}{|\mathcal{V}_m|}} \text{ is finite and positive.} \tag{4.8}$$

Then,

$$\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}\left(\hat\beta-\beta_0\right)=\left(1+\lambda^{-2}\right)A\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}S_{n,m}+o_P(1), \tag{4.9}$$

as $n,m\to\infty$, where

$$S_{n,m}=\mathbf{s}_{\circ n}-\lambda^2B\mathbf{s}_{\vee m},$$
$$\mathbf{s}_{\circ n}=\sum_{t\in\mathcal{P}_n}\mathbf{z}_{\circ t}\left(y_{\circ t}-\mathbf{z}_{\circ t}^\top\left[\mathbf{E}\left(\mathbf{z}_\circ\mathbf{z}_\circ^\top\right)\right]^{-1}\mathbf{E}\left(\mathbf{z}_\circ y_\circ\right)\right),$$
$$\mathbf{s}_{\vee m}=\sum_{t\in\mathcal{V}_m}\mathbf{z}_{\vee t}\left(g(\mathbf{x}_{\vee t},\beta_0)-\mathbf{z}_{\vee t}^\top\left[\mathbf{E}\left(\mathbf{z}_\vee\mathbf{z}_\vee^\top\right)\right]^{-1}\mathbf{E}\left(\mathbf{z}_\vee g\left(\mathbf{x}_\vee,\beta_0\right)\right)\right),$$
$$A=\left\{\mathbf{E}\left[\frac{\partial g\left(\mathbf{x}_\vee,\beta_0\right)}{\partial\beta}\mathbf{z}_\vee^\top\right]\left[\mathbf{E}\left(\mathbf{z}_\vee\mathbf{z}_\vee^\top\right)\right]^{-1}\mathbf{E}\left(\mathbf{z}_\circ\mathbf{z}_\circ^\top\right)\left[\mathbf{E}\left(\mathbf{z}_\vee\mathbf{z}_\vee^\top\right)\right]^{-1}\mathbf{E}\left[\mathbf{z}_\vee\frac{\partial g(\mathbf{x}_\vee,\beta_0)}{\partial\beta^\top}\right]\right\}^{-1}$$
$$\times\left\{\mathbf{E}\left[\frac{\partial g(\mathbf{x}_\vee,\beta_0)}{\partial\beta}\mathbf{z}_\vee^\top\right]\left[\mathbf{E}\left(\mathbf{z}_\vee\mathbf{z}_\vee^\top\right)\right]^{-1}\right\},$$
$$B=\mathbf{E}\left(\mathbf{z}_\circ\mathbf{z}_\circ^\top\right)\left[\mathbf{E}\left(\mathbf{z}_\vee\mathbf{z}_\vee^\top\right)\right]^{-1}.$$

PROOF. We apply Lemma 4.1 and the same Taylor series expansion as in [1]. The estimator $\hat\beta$ satisfies the first-order condition

$$\frac{\partial Q_{n,m}\left(\hat\beta\right)}{\partial\beta}=c_n\frac{\partial g^\top\left(\mathbf{X}_\vee,\hat\beta\right)}{\partial\beta}\mathbf{Z}_\vee\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\mathbf{Z}_\circ^\top\left(\mathbf{Y}_\circ-\mathbf{Z}_\circ\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\mathbf{Z}_\vee^\top g\left(\mathbf{X}_\vee,\hat\beta\right)\right)=\mathbf{0},$$

where $c_n=-2/|\mathcal{P}_n|$. By a Taylor series expansion

$$\begin{aligned}\mathbf{0}={}&\frac{\partial g^\top\left(\mathbf{X}_\vee,\beta_0\right)}{\partial\beta}\mathbf{Z}_\vee\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\mathbf{Z}_\circ^\top\left(\mathbf{Y}_\circ-\mathbf{Z}_\circ\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\mathbf{Z}_\vee^\top g\left(\mathbf{X}_\vee,\beta_0\right)\right)\\&+\left\{\left[\frac{\partial^2 g^\top\left(\mathbf{X}_\vee,\overline\beta\right)}{\partial\beta\partial\beta_1}\mathbf{Z}_\vee,\dots,\frac{\partial^2 g^\top\left(\mathbf{X}_\vee,\overline\beta\right)}{\partial\beta\partial\beta_l}\mathbf{Z}_\vee\right]\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\right.\\&\times\mathbf{Z}_\circ^\top\left(\mathbf{Y}_\circ-\mathbf{Z}_\circ\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\mathbf{Z}_\vee^\top g\left(\mathbf{X}_\vee,\overline\beta\right)\right)\\&\left.-\left(\frac{\partial g^\top\left(\mathbf{X}_\vee,\overline\beta\right)}{\partial\beta}\mathbf{Z}_\vee\right)\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\left(\mathbf{Z}_\circ^\top\mathbf{Z}_\circ\right)\left(\mathbf{Z}_\vee^\top\mathbf{Z}_\vee\right)^{-1}\left(\mathbf{Z}_\vee^\top\frac{\partial g\left(\mathbf{X}_\vee,\overline\beta\right)}{\partial\beta^\top}\right)\right\}\left(\hat\beta-\beta_0\right),\end{aligned}\tag{4.10}$$

where the coordinates of $\overline{\beta}$ are between those of β_0 and $\hat{\beta}$. Using (A2′), (4.3), (4.5), and (4.6) we obtain that $\mathbf{E}[g(\mathbf{x}_{\vee},\beta)\mathbf{z}_{\vee}^{\top}]$, $\mathbf{E}[\frac{\partial g(\mathbf{x}_{\vee},\beta)}{\partial\beta}\mathbf{z}_{\vee}^{\top}]$,and $\mathbf{E}[\frac{\partial^2 g(\mathbf{x}_{\vee},\beta)}{\partial\beta\partial\beta_j}\mathbf{z}_{\vee}^{\top}]$ are continuous in β. Therefore, by the ULLN, $1/|\mathcal{V}_m|\frac{\partial g^{\top}(\mathbf{X}_{\vee},\beta_0)}{\partial\beta}\mathbf{Z}_{\vee}$, $1/|\mathcal{V}_m|\frac{\partial g^{\top}(\mathbf{X}_{\vee},\overline{\beta})}{\partial\beta}\mathbf{Z}_{\vee}$, $1/|\mathcal{V}_m|\frac{\partial^2 g^{\top}(\mathbf{X}_{\vee},\overline{\beta})}{\partial\beta\partial\beta_j}\mathbf{Z}_{\vee}$ $(j=1,2,\ldots,l)$ converge in probability to $\mathbf{E}[\frac{\partial g(\mathbf{x}_{\vee},\beta_0)}{\partial\beta}\mathbf{z}_{\vee}^{\top}]$, $\mathbf{E}[\frac{\partial g(\mathbf{x}_{\vee},\beta_0)}{\partial\beta}\mathbf{z}_{\vee}^{\top}]$, $\mathbf{E}[\frac{\partial^2 g(\mathbf{x}_{\vee},\beta_0)}{\partial\beta\partial\beta_j}\mathbf{z}_{\vee}^{\top}]$ $(j=1,2,\ldots,l)$, respectively, as $n,m\to\infty$. Moreover,

$$\frac{1}{|\mathcal{P}_n|}\mathbf{Z}_{\circ}^{\top}\left(\mathbf{Y}_{\circ}-\mathbf{Z}_{\circ}\left(\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\right)^{-1}\mathbf{Z}_{\vee}^{\top}g\left(\mathbf{X}_{\vee},\overline{\beta}\right)\right)$$
$$\to\mathbf{E}\left(\mathbf{z}_{\circ}y_{\circ}\right)-\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\left[\mathbf{E}\left(\mathbf{z}_{\vee}\mathbf{z}_{\vee}^{\top}\right)\right]^{-1}\mathbf{E}\left(\mathbf{z}_{\vee}g\left(\mathbf{x}_{\vee},\beta_0\right)\right)=0.$$

The last equality is the consequence of (A7). Multiplying (4.10) with $\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}$, using the above limit relations, and applying condition (4.7), we obtain,

$$\begin{aligned}\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}\left(\hat{\beta}-\beta_0\right)&=\left\{\mathbf{E}\left[\frac{\partial g\left(\mathbf{x}_{\vee},\beta_0\right)}{\partial\beta}\mathbf{z}_{\vee}^{\top}\right]\left[\mathbf{E}\left(\mathbf{z}_{\vee}\mathbf{z}_{\vee}^{\top}\right)\right]^{-1}\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\right.\\&\quad\left.\times\left[\mathbf{E}\left(\mathbf{z}_{\vee}\mathbf{z}_{\vee}^{\top}\right)\right]^{-1}\mathbf{E}\left[\mathbf{z}_{\vee}\frac{\partial g\left(\mathbf{x}_{\vee},\beta_0\right)}{\partial\beta^{\top}}\right]+o_P(1)\right\}^{-1}\\&\quad\times\left\{\mathbf{E}\left[\frac{\partial g\left(\mathbf{x}_{\vee},\beta_0\right)}{\partial\beta}\mathbf{z}_{\vee}^{\top}\right]\left[\mathbf{E}\left(\mathbf{z}_{\vee}\mathbf{z}_{\vee}^{\top}\right)\right]^{-1}+o_P(1)\right\}\\&\times\frac{1+\lambda^{-2}}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\left\{\mathbf{Z}_{\circ}^{\top}\left(\mathbf{Y}_{\circ}-\mathbf{Z}_{\circ}\left(\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\right)^{-1}\mathbf{Z}_{\vee}^{\top}g\left(\mathbf{X}_{\vee},\beta_0\right)\right)+o_P(1)\right\}.\end{aligned}\tag{4.11}$$

Observe that for the above step the boundedness (in probability) of

$$\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}S_{n,m}=\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\mathbf{Z}_{\circ}^{\top}\left(\mathbf{Y}_{\circ}-\mathbf{Z}_{\circ}\left(\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\right)^{-1}\mathbf{Z}_{\vee}^{\top}g(\mathbf{X}_{\vee},\beta_0)\right)\tag{4.12}$$

was necessary. However, this will be implied by the argument below. To separate primary and validation data values, consider the following:

$$\begin{aligned}\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}S_{n,m}&=\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\mathbf{Z}_{\circ}^{\top}\left(\mathbf{Y}_{\circ}-\mathbf{Z}_{\circ}\left[\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\right]^{-1}\mathbf{E}(\mathbf{z}_{\circ}y_{\circ})\right)\\&\quad-\left(\frac{|\mathcal{P}_n|}{|\mathcal{V}_m|}\right)\left(\frac{1}{|\mathcal{P}_n|}\mathbf{Z}_{\circ}^{\top}\mathbf{Z}_{\circ}\right)\left[\frac{1}{|\mathcal{V}_m|}\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\right]^{-1}\\&\times\left\{\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\left(\mathbf{Z}_{\vee}^{\top}g(\mathbf{X}_{\vee},\beta_0)-\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\left[\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\right]^{-1}\mathbf{E}(\mathbf{z}_{\circ}y_{\circ})\right)\right\}\\&=\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\mathbf{Z}_{\circ}^{\top}\left(\mathbf{Y}_{\circ}-\mathbf{Z}_{\circ}\left[\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\right]^{-1}\mathbf{E}(\mathbf{z}_{\circ}y_{\circ})\right)-\lambda^2\mathbf{E}(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top})\left[\mathbf{E}(\mathbf{z}_{\vee}\mathbf{z}_{\vee}^{\top})\right]^{-1}\\&\quad\times\left\{\frac{1}{\sqrt{|\mathcal{P}_n|+|\mathcal{V}_m|}}\left(\mathbf{Z}_{\vee}^{\top}g(\mathbf{X}_{\vee},\beta_0)-\mathbf{Z}_{\vee}^{\top}\mathbf{Z}_{\vee}\left[\mathbf{E}\left(\mathbf{z}_{\circ}\mathbf{z}_{\circ}^{\top}\right)\right]^{-1}\mathbf{E}(\mathbf{z}_{\circ}y_{\circ})\right)\right\}+o_P(1)\end{aligned}\tag{4.13}$$

because $\lim_{n,m\to\infty}\sqrt{|\mathcal{P}_n|/|\mathcal{V}_m|}=\lambda$, the law of large numbers, and that the expression above in $\{\ \}$ is bounded in probability. The latest fact follows from (A5), (A6′) and the CLT for random fields [8, Theorem 3.3.1]. Now, (4.11) and (4.13), and the fact that summands in (4.13) are bounded in probability imply the result. ∎

REMARK. One can obtain the same expression for $\sqrt{|\mathcal{P}_n|}(\hat{\beta}-\beta_0)$ as in Proposition 2.2 of [1]. However, that form is suitable to show asymptotic normality when primary and validation data are independent.

Now, we formulate the asymptotic normality result.

THEOREM 4.4. *In addition to assumptions of Theorem 4.3, suppose that condition (4.1) and*

$$\sum_{r=1}^{\infty} r^{d-1}\alpha_\eta(r;u,v)] < \infty, \qquad \text{if } u+v \le 4 \text{ and } \alpha_\eta(r;1,\infty) = o\left(r^{-d}\right),$$

are satisfied for $\eta_t = (y_{\circ t}, \mathbf{z}_{\circ t}, g(\mathbf{x}_{\vee t},\beta), \mathbf{z}_{\vee t})$. *Moreover, assume that*

$$\liminf_{n,m\to\infty} \lambda_{\min}\left((|\mathcal{P}_n| + |\mathcal{V}_m|)^{-1}\Sigma_{n,m}\right) > 0,$$

where $\Sigma_{n,m}$ *denotes the covariance matrix of* $S_{n,m}$ *and* $\lambda_{\min}$ *denotes the smallest eigenvalue. Then,* $\sqrt{|\mathcal{P}_n| + |\mathcal{V}_m|}\,\Sigma_{n,m}^{-1/2}(\hat\beta - \beta_0)$ *is asymptotically normally distributed with zero mean, as* $n, m \to \infty$.

PROOF. This is a consequence of Theorem 4.3 and the CLT for random fields, see [8, Theorem 3.3.1]. ∎

5. INFILL ASYMPTOTICS FOR SPATIAL MODELS

In this section, we shall prove that the Lee-Sepanski estimator is not consistent if the observations become dense in a fixed domain. To this end, we find the limit $Q_\infty(\beta)$ of $Q_{n,m}(\beta)$, then we find conditions ensuring β_0 not to be the minimum point of $Q_\infty(\beta)$.

Let $T \subset \mathbf{R}^d$ be a compact domain and

$$y_t = g(\mathbf{x}_t, \beta_0) + \varepsilon_t, \tag{5.1}$$

$t \in T$, be a random field. Explanatory variables $\mathbf{x}_t$ are measured with error

$$\tilde{\mathbf{x}}_t \text{ measures } \mathbf{x}_t \quad \text{with error.} \tag{5.2}$$

Let $T_\circ \subseteq T$ be a fixed compact domain. Primary data are observations of the random field on $T_\circ$. Now, we describe the infill sampling design, similar to the one in [9]. Let $\mathcal{P}_n$, $n = 1, 2, \dots$, be a sequence of finite subsets of $T_\circ$ satisfying the following conditions. Cover $T_\circ$ with congruent d-dimensional rectangles. Choose one point from the intersection of $T_\circ$ and each rectangle (if the intersection is not empty). Then $\mathcal{P}_n$ consists of these points. When we study asymptotic behaviour, we suppose that the edges of the covering rectangles converge to 0, as $n \to \infty$. $\mathcal{P}_n$ is the n^{th} set of points, where the random field is observed. The primary data are $(y_{\circ t}, \tilde{\mathbf{x}}_{\circ t})$, $t \in \mathcal{P}_n$. Let $\mathbf{Y}_\circ$ and $\mathbf{Z}_\circ$ be defined as in Section 2. These quantities are based on the primary data.

Assume that validation data are available: $(\mathbf{x}_{\vee t}, \tilde{\mathbf{x}}_{\vee t})$, $t \in \mathcal{V}_m$. Here $\mathcal{V}_m \subset T_\vee$, $m = 1, 2, \dots$, where $T_\vee \subseteq T$ is a fixed compact domain. Assume that the sequence $\mathcal{V}_m$, $m = 1, 2, \dots$, satisfies the same type of conditions as $\mathcal{P}_n$, $n = 1, 2, \dots$. Let $g(\mathbf{X}_\vee, \beta), \mathbf{Z}_\vee$ be defined as in Section 2. These quantities are based on the validation data.

The estimator $\hat\beta$ is the minimum point of

$$Q_{n,m}(\beta) = \frac{\mathrm{vol}(T_\circ)}{|\mathcal{P}_n|} \left\|\mathbf{Y}_\circ - \mathbf{Z}_\circ(\mathbf{Z}_\vee^\top \mathbf{Z}_\vee)^{-1}\mathbf{Z}_\vee^\top g(\mathbf{X}_\vee, \beta)\right\|^2, \tag{5.3}$$

where $\mathrm{vol}(T_\circ)$ is the volume of $T_\circ$, while $|\mathcal{P}_n|$ is the cardinality of $\mathcal{P}_n$.

ASSUMPTIONS.

(B3) For the primary data: $y_{\circ t}$ and $\mathbf{z}_{\circ t}$, $t \in T_\circ$, are L_4-continuous random fields. $\int_{T_\circ} \mathbf{z}_{\circ t}\mathbf{z}_{\circ t}^\top\, dt$ is nonsingular.

(B4) For the validation data: $g(\mathbf{x}_{\vee t}, \beta)$ and $\mathbf{z}_{\vee t}$, $t \in T_\vee$, are L_4-continuous random fields. $\int_{T_\vee} \mathbf{z}_{\vee t}\mathbf{z}_{\vee t}^\top\, dt$ is nonsingular.

(B5)

$$\sup_{t\in T_\vee}\sup_{\beta\in\Theta} \mathbf{E}\left\|\frac{\partial g(\mathbf{x}_{\vee t},\beta)}{\partial\beta}\right\|^4 < \infty.$$

Now, let

$$Q_\infty(\beta) = a_1 + a_2, \tag{5.4}$$

where

$$a_1 = \left[\left(\int_{T_\circ} {y_{\circ t}}^2\,dt\right) - \left(\int_{T_\circ} y_{\circ t}\mathbf{z}_{\circ t}^\top\,dt\right)\left(\int_{T_\circ} \mathbf{z}_{\circ t}\mathbf{z}_{\circ t}^\top\,dt\right)^{-1}\left(\int_{T_\circ} \mathbf{z}_{\circ t}y_{\circ t}\,dt\right)\right], \tag{5.5}$$

$$a_2 = \mathbf{b}^\top\mathbf{A}\mathbf{b}, \qquad \mathbf{A} = \left[\int_{T_\circ} \mathbf{z}_{\circ t}\mathbf{z}_{\circ t}^\top\,dt\right], \tag{5.6}$$

$$\mathbf{b}^\top = \left[\left(\int_{T_\circ} y_{\circ t}\mathbf{z}_{\circ t}^\top\,dt\right)\left(\int_{T_\circ} \mathbf{z}_{\circ t}\mathbf{z}_{\circ t}^\top\,dt\right)^{-1} - \left(\int_{T_\vee} g(\mathbf{x}_{\vee t},\beta)\mathbf{z}_{\vee t}^\top\,dt\right)\left(\int_{T_\vee} \mathbf{z}_{\vee t}\mathbf{z}_{\vee t}^\top\,dt\right)^{-1}\right], \tag{5.7}$$

where the integrals are meant in L_2 sense.

THEOREM 5.1. *Assume that (A1), (A2), (B3)–(B5) are satisfied. Then $Q_{n,m}(\beta) \to Q_\infty(\beta)$ in probability, uniformly in β, as $n,m \to \infty$.*

PROOF. The definition of the L_2-integral and (B3), (B4) imply that partial sums in (4.4) converge to the corresponding integrals in L_2. Moreover, assumption (B5) implies that integrals depending on β converge uniformly in β. Therefore, $Q_{n,m}(\beta) \to Q_\infty(\beta)$ in probability, uniformly in β, as n, $m \to \infty$. ∎

THEOREM 5.2.

(a) *Assume that (A1), (A2), (B3)–(B5) are satisfied. Suppose that $T_\circ = T_\vee$, moreover, $(y_{\circ t}, \tilde{\mathbf{x}}_{\circ t}) = (y_t, \tilde{\mathbf{x}}_t)$ and $(\mathbf{x}_{\vee t}, \tilde{\mathbf{x}}_{\vee t}) = (\mathbf{x}_t, \tilde{\mathbf{x}}_t)$ for each t. Let $\mathbf{A}$ and Q_∞ be defined in (5.6) and (5.4), respectively. Let*

$$f_1 = \max\left\{\int |(\mathbf{z}_t)_h| \left|\frac{\partial g(\mathbf{x}_t,\beta_0)}{\partial\beta_i}\right| dt : i = 1,\dots,l,\ h = 1,\dots,k_1\right\},$$

$$f_2 = \max\left\{\int |(\mathbf{z}_t)_h| \sup_{\beta\in\Theta}\left|\frac{\partial^2 g(\mathbf{x}_t,\beta)}{\partial\beta_i\partial\beta_j}\right| dt : i,j = 1,\dots,l,\ h = 1,\dots,k_1\right\},$$

$$f_3 = \|\mathbf{A}^{-1}\|,$$

where every integral is taken upon $T_\circ = T_\vee$ (and subscript h stands for the coordinate of a vector). Suppose that f_1, f_2, f_3 are finite (not necessarily uniformly on Ω). Let

$$\left(\int \frac{\partial g(\mathbf{x}_t,\beta_0)}{\partial\beta}\mathbf{z}_t^\top\,dt\right)\left(\int \mathbf{z}_t\mathbf{z}_t^\top\,dt\right)^{-1}\left(\int \mathbf{z}_t\varepsilon_t\,dt\right) \neq 0, \tag{5.8}$$

almost surely. Then, β_0 is not a minimum point of $Q_\infty(\beta)$, a.s.

(b) *If, moreover, $Q_\infty(\beta)$ is a.s. continuous in β, then $\hat\beta$ is not consistent.*

PROOF. Expand $g(\mathbf{x}_t,\beta)$ around β_0 to Taylor series up to second order. ∎

REFERENCES

1. L.-F. Lee and J.H. Sepanski, Estimation of linear and nonlinear errors-in-variables models using validation data, *J. Amer. Statist. Assoc.* **90**, 130–140 (1995).
2. W.A. Fuller, *Measurement Error Models*, Wiley, New York, (1987).
3. I. Fazekas, J. Lauridsen and S. Baran, Asymptotic properties in space and time of an estimator in errors-in-variables models in the presence of validation data, *Occasional Papers*, Odense University (3) (1997).
4. MATLAB Reference Guide, The MathWorks, Inc. (1992).
5. P. Doukhan, Mixing. Properties and examples, *Lecture Notes in Statistics*, Springer, New York **85** (1994).
6. T. Amemiya, *Advanced Econometrics*, Basil Blackwell, Oxford, (1985).
7. I. Fazekas and A. Kukush, Asymptotic properties of an estimator in nonlinear functional errors-in-variables models with dependent error terms, *Computers Math. Appl.* **34** (10), 23–39 (1997).
8. X. Guyon, *Random Fields on a Network. Modeling, Statistics, and Applications*, Springer-Verlag, New York, (1995).
9. S.N. Lahiri, On inconsistency of estimators based on spatial data under infill asymptotics, *Sankhya* **58 Ser. A**, 403–417 (1996).